NOT TO BE CITED WITHOUT PRIOR
REFERENCE TO THE AUTHOR(S)

Northwest Atlantic Fisheries Organization

Serial No. N6062

NAFO SCR Doc. 12/035

SCIENTIFIC COUNCIL MEETING – JUNE 2012

A stochastic VPA of the Flemish Cap cod stock

by

Antonio Vázquez and Mónica Mandado

Instituto de Investigaciones Marinas
Muelle de Bouzas, Vigo, Spain

Abstract

An ADAPT of the cod stock in Flemish Cap is repeated with alternative input data produced by bootstrap. A new set of survey indices was calculated in each iteration by re-sampling the original detailed survey data. Results are compared with a Bayesian solution. Even results are different, no main differences are observed in the distribution of parameters, but Bayesians exhibit higher skewness. This analysis is not proposed as an alternative assessment of the 3M cod stock, but to check Bayesian results.

KEYWORDS: Stochastic VPA, Flemish Cap, Cod.

Introduction

The main benefit of stochastic modelling as compared to deterministic virtual population analysis is that, instead of point estimates, posterior distributions of plausible parameter values are obtained. A stochastic approach can be solved analytically (e.g. Lewy 1988) or by re-sampling techniques: Bayesian statistics and bootstrap being the most common techniques. The last assessment of the cod stock in NAFO Division 3M (González-Troncoso and Vázquez 2011) is based on a VPA solved by a Bayesian approach (Fernández *et al.*, 2007, 2008). In our view, a bootstrap analysis could capture the variability of VPA input data better than the Bayesian one: VPA input data produced by bootstrap contains roughly their real variance–covariance, so no assumption is required on their distribution type and deviation magnitude. In order to make a comparison of results of both methods we developed a bootstrapped VPA to compare it with the Bayesian assessment results for 3M cod, making use of the detailed original survey data.

Material and methods

The EU survey is carried out annually in July, and results are regularly published (Vazquez 2012). Table 1 shows the survey indices of abundance at age. These data are not exactly equal the ones presented by González-Troncoso and Vázquez (2011) because a different algorithm was used to calculate length frequency distribution: it was calculated by raising total sample distribution to the estimated biomass, and now it was done by raising mean frequency per mile from mean sample weight per mile to the estimated biomass.

The VPA was solved by ADAPT (Gavaris 1988) with the following peculiarities:

– Age 8 was a plus group, so it was excluded from the fit.

- – Every cohort has a parameter, which is the survivor from its last age or year. It implied 28 parameters (N(age,2011), age=2…8, and N(8,year), year=1990…2010). Good behaviour of this formulation has been recently checked (Vázquez *et al*. 2009).
- – In 2002-2005 there was not catch at age in numbers, but only total catch in weight. Fishing mortalities in those years are set in such a way that they would produce the same catches as reported landings. Catch at age 8+ was equal to catch at ages 5-7 raised in the same proportion as survey results at those ages.
- – The 8+ group abundance is calculated after the VPA is finished and it is based on the reported catch, assuming fishing mortality at age 8+ equal F at age 7. The result is compared with the parameter for abundance at age 8 in that year and the highest is chosen.
- – F_{bar} is calculated as a mean for ages 3 to 5, as it was in the Bayesian run.

Bootstrap on survey indices at age was done in three levels (Cerviño and Vázquez 2001):

1- Hauls in each stratum were re-sampled with replacement, keeping the original number of hauls in each stratum.
2- Length distribution of each selected haul was re-sampled with replacement, keeping the original number of measurements.
3- Age-length-key was rebuilt by re-sampling with replacement ages in each length class, keeping the same number of identified ages in each of them.

A deterministic initial ADAPT run with the original survey data provides the reference results for distribution of estimated parameters of later bootstrap. Distributions presented in Figures 3, 4 and 5 are normalized, dividing every parameter by its deterministic value, so all distributions are around the 1.0 value.

Catch at age 1988-2010 is presented in Table 2; it is the same used in the assessment.

Results

Both formulations are quite similar. Main common feature is the variability on input data for survey abundance indices: it is produced by bootstrap in the ADAPT formulation and by priors in the Bayesian approach.

Main differences are in the way fishing mortality is calculated in years 2002-2005, where catch at age data is not available. The Bayesian formulation introduces a prior for each age and year, which means 28 additional parameters. No new parameters are introduced for ADAPT because annual F is deterministically calculated as to produce the specify catch, assuming partial recruitment and mean weight at age equal those in the following year.

A minor difference is on Natural Mortality (M). It was agreed that be calculated with uncertainty in the assessment Bayesian formulation, so a prior over this indices is applied since then (González-Troncoso and Fernández 2009). A fix 0.2 value is used for ADAPT. Calculating M rises the Bayesian number of parameters to 58 while ADAPT only needs 28.

Abundance at age for 1988-2010 from the ADAPT initial deterministic analysis is presented in Table 3. The table includes an additional column for age 8+ and an additional file for 2011 where assessment (Bayesian) results are presented; differences in survivors are quite obvious. The relative similarity between these two estimates for age 8+ abundance is low (r^2 = 0.38) but it is better for last year survivors ages 2 to 7 (r^2 = 0.92) (r^2 = 0.98 if age 2 is excluded). Differences in age 8+ should be attributed to differences in the method to calculate them; those figures are calculated once the VPA is completed in both cases. Differences in annual fishing mortality and SSB (Figure 1) show same trends in spite of differences. Figure 2 shows the relationship between abundances of all years and at all ages in the assessment and the deterministic ADAPT; agreement is quite good in most cases; main differences occur in last years.

Distribution of survivors' abundance from bootstrap is presented in Figure 3. It is observed that results of most recent years are better defined that initial years, but there is not a common pattern: well defined distributed years alternate with the badly defined ones.

A comparison of the Bayesian posterior distribution of last year survivors and the distribution of the same parameters from bootstrap is presented in Figure 4; relative width distributions of the same variable seem to be similar.

Figure 5 shows distributions of SSB, fishable biomass and 5+ biomass in survivors from bootstrap results. In order to compare them with the Bayesian results only the 5% and 95% percentiles are available for SSB, and they are indicated in the same Figure 5. These percentile limits indicate more asymmetric distribution of SSB results in the Bayesian analysis than in the bootstrap procedure, which means higher positive skewness.

Discussion

This bootstrap analysis only includes variability on survey results. This is possible because survey data are stored and processed in a standard way, having access to all data at the same time. If commercial catch sampling data was available in the same way it would be possible to include variability of catch at age estimates in the analysis. Then, results would be calculated in a full stochastic way.

Distributions of last year survivors from both Bayesian and bootstrap analyses have similar deviation (variance). This similarity could be understood as both methods being robust; however variability in bootstrap results is produced by change in input data and it is not affected by goodness of fit, while variability in the Bayesian analysis comes from goodness of fit and only slightly from priors on input data. We concluded this similarity is spurious, and new formulations are required to verify it.

References

Cerviño, S. and A. Vázquez – 2001. Variability and abundance indices and its progression through age-structured models: an stochastic simulation with Flemish Cap cod. *NAFO SCR Doc.* 01/56.

Fernández, C., S. Cerviño and A. Vázquez – 2007. A Survey-based assessment of cod in division 3M. NAFO SCR Doc. 07/39.

Fernández, C., S. Cerviño, F. Saborido-Rey and A. Vázquez – 2008. Assessment of the Cod Stock in NAFO Division 3M. NAFO SCR Doc. 08/26.

Gavaris, S.– 1988. An adaptive framework for the estimation of population size. *Can. Atl. Fish. Adv. Comm. (CAFSAC) Res. Doc.* 88/29.

González-Troncoso, D. and C. Fernández – 2009. Assessment of the Cod Stock in NAFO Division 3M. *NAFO SCR Doc.* 09/034.

González-Troncoso, D. and A. Vázquez – 2011. Assessment of the Cod Stock in NAFO Division 3M. *NAFO SCR Doc*. 11/38.

Lewy P.– 1988. Integrated stochastic virtual population analysis: estimates and their precision of fishing mortalities and stock sizes for the North Sea whiting stock. *J. Cons. int. Explor. Mer* 44 (3): 217-228.

Vázquez, A., R. Piñeiro and M. Mandado – 2009. Inquiry into ADAPT's performance. *NAFO SCR Doc*. 09/049.

Vázquez, A.– 2012. Results from Bottom Trawl Survey on Flemish Cap of July 2011. *NAFO SCR Doc.* 12/26.

Table 1 – Survey indices of abundance at age 1988-2010.

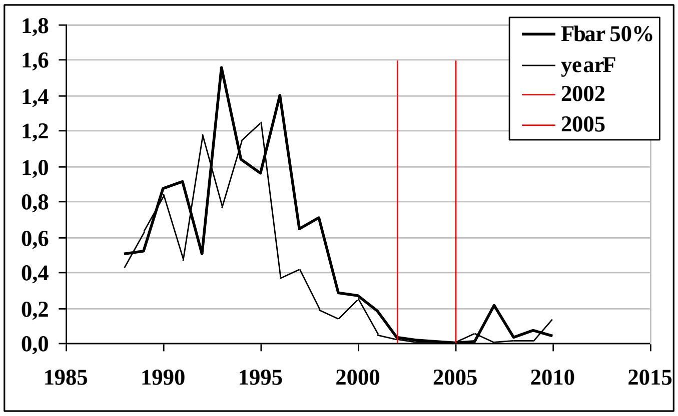
	age							
year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8+**
1988	4315	71506	40874	10740	1154	182	211	67
1989	21479	11234	87848	51078	19331	1307	161	149
1990	2506	11837	4677	15613	14674	4221	346	279
1991	142064	26013	16757	2026	6350	1769	310	136
1992	71768	37874	5215	1998	321	1248	232	20
1993	4273	133116	27930	989	1234	164	485	98
1994	3169	3780	24835	4583	122	65	7	122
1995	1538	11357	1231	3571	886	33	25	25
1996	37	2930	6164	822	2270	189	8	6
1997	40	140	3162	4409	360	912	20	0
1998	24	76	85	1136	1444	72	143	7
1999	6	78	103	105	657	416	19	6
2000	173	13	271	170	83	401	159	28
2001	457	1645	12	103	70	13	134	116
2002	0	1070	489	20	57	31	26	130
2003	613	50	612	132	22	47	7	71
2004	0	3480	41	623	169	5	11	24
2005	8142	16	1002	68	640	129	0	31
2006	19656	3926	62	1475	85	593	116	27
2007	3898	11868	5052	21	1138	58	424	106
2008	6055	16731	12450	4537	72	957	57	326
2009	5133	7602	16392	14368	4171	26	1098	339
2010	66827	27456	8750	7687	4977	1790	8	816

Table 2 – Catch in numbers (thousands) at age 1988-2010.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8+**
1988	1	3500	25593	11161	1399	414	315	162
1989	0	52	15399	23233	9373	943	220	205
1990	7	254	2180	15740	10824	2286	378	117
1991	1	561	5196	1960	3151	1688	368	76
1992	0	15517	10180	4865	3399	2483	1106	472
1993	0	2657	14530	3547	931	284	426	213
1994	0	1219	25400	8273	386	185	14	182
1995	0	0	264	6553	2750	651	135	232
1996	0	81	714	311	1072	88	0	0
1997	0	0	810	762	143	286	48	0
1998	0	0	8	170	286	30	19	2
1999	0	0	15	15	96	60	3	1
2000	0	10	54	1	1	4	1	0
2001	0	9	0	4	2	0	2	2
2002								
2003								
2004								
2005								
2006	0	22	19	81	2	10	2	0
2007	0	2	30	1	27	1	14	5
2008	1	89	136	133	3	40	1	3
2009	0	23	51	210	108	0	32	7
2010	34	452	1145	1498	808	388	4	103

Table 3 – Abundance (thousands) at age for 1988-2010 from the ADAPT initial analysis. Additional highlighted figures for age 8+ and year 2011, are the results in the assessment document (González-Troncoso and Vázquez 2011). Years without catch at age information (2002-2005) are also marked.

year	1	2	3	4	5	6	7	8+	8+
1988	16547	65805	83025	31282	4134	946	883	454	240
1989	21732	13547	50718	45014	15611	2131	404	377	317
1990	27077	17793	11044	27707	16149	4459	902	279	190
1991	67834	22163	14338	7081	8687	3645	1613	400	130
1992	62981	55537	17639	7084	4037	4290	1477	990	538
1993	3393	51564	31538	5392	1500	340	1305	653	329
1994	6309	2778	39819	12846	1274	402	30	687	678
1995	3640	5166	1185	10088	3187	697	164	282	317
1996	169	2981	4229	733	2455	216	12	16	1
1997	154	139	2367	2820	322	1052	98	10	1
1998	226	126	113	1212	1624	136	604	64	27
1999	37	185	104	86	839	1072	84	478	25
2000	653	30	151	71	57	601	824	66	1
2001	1623	535	16	75	57	45	488	674	162
2002	160	1329	430	13	58	45	37	398	275
2003	3393	131	1086	342	10	47	37	30	267
2004	106	2778	107	881	278	8	38	30	265
2005	9452	87	2273	86	716	227	7	31	267
2006	22631	7739	71	1850	71	584	185	6	23
2007	18171	18529	6316	41	1442	56	469	168	74
2008	13506	14877	15168	5144	33	1156	45	372	66
2009	22843	11057	12100	12296	4092	24	910	199	88
2010	177535	18703	9032	9861	9877	3252	20	716	473
2011		145322	14904	6363	6724	7358	2313	13	324
2011		22697	9132	3720	3551	3140	1025	324	



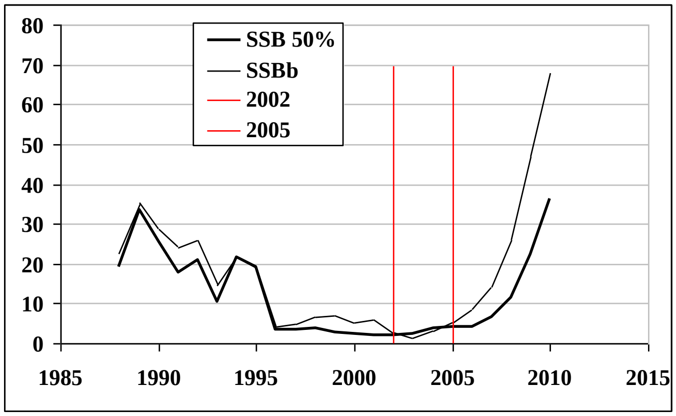
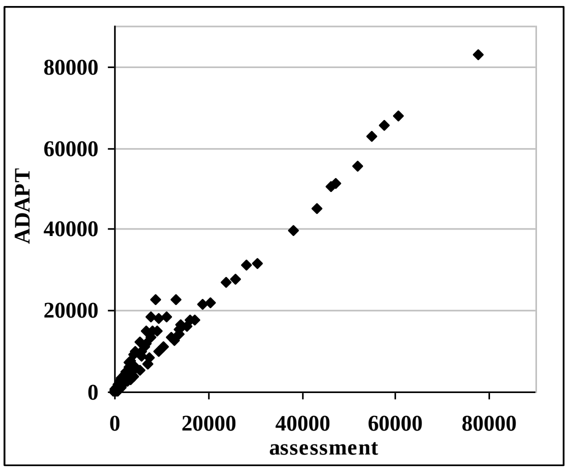


Figure 1 – Annual fishing mortality (up) and SSB (down, thousand tons) as calculated by the Bayesian (Fbar 50%, SSB 50%) and bootstrap (yearF, SSBb) methods.



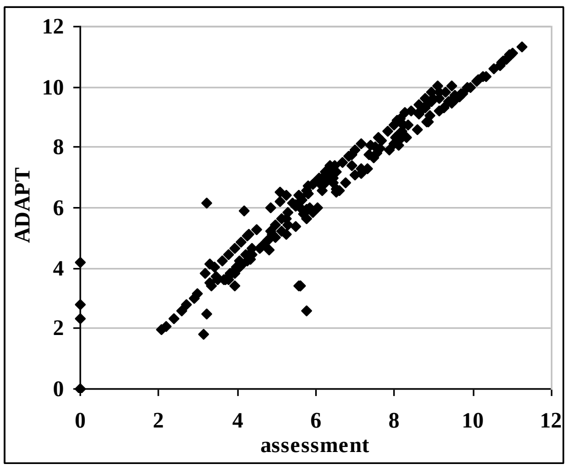


Figure 2 – Abundance at age in all years as calculated by ADAPT vs the assessment results. Left: absolute numbers, right: logarithms.

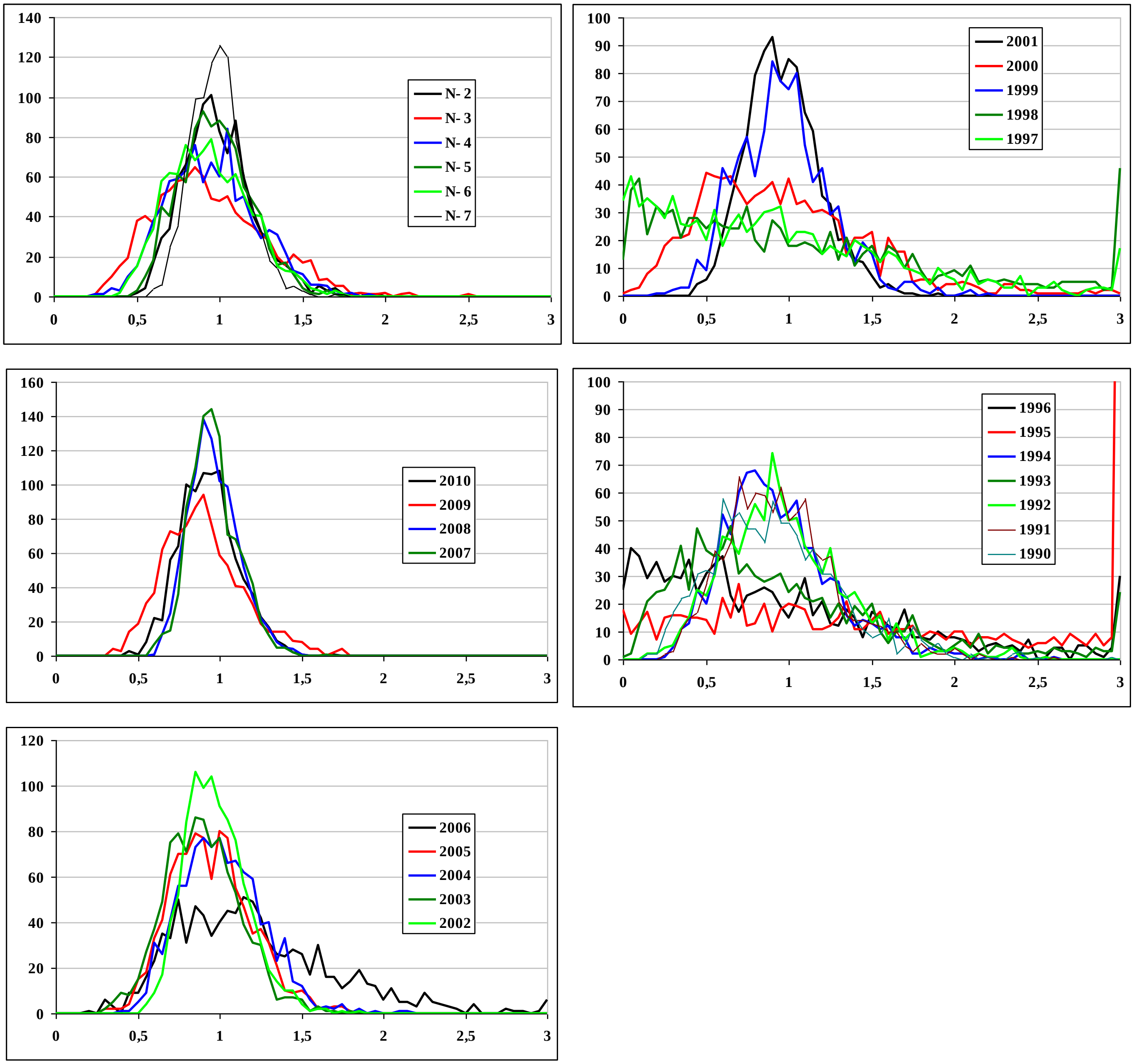


Figure 3 – Distribution of survivors' abundance from bootstrap. Survivors in 2011 are at the indicated ages in left square. Survivors at age 8 and the indicated year are in all others squares. All abundances were normalized: divided by the corresponding figure in the initial deterministic run to unify the scale in the X axis. The Y axis indicates frequency.

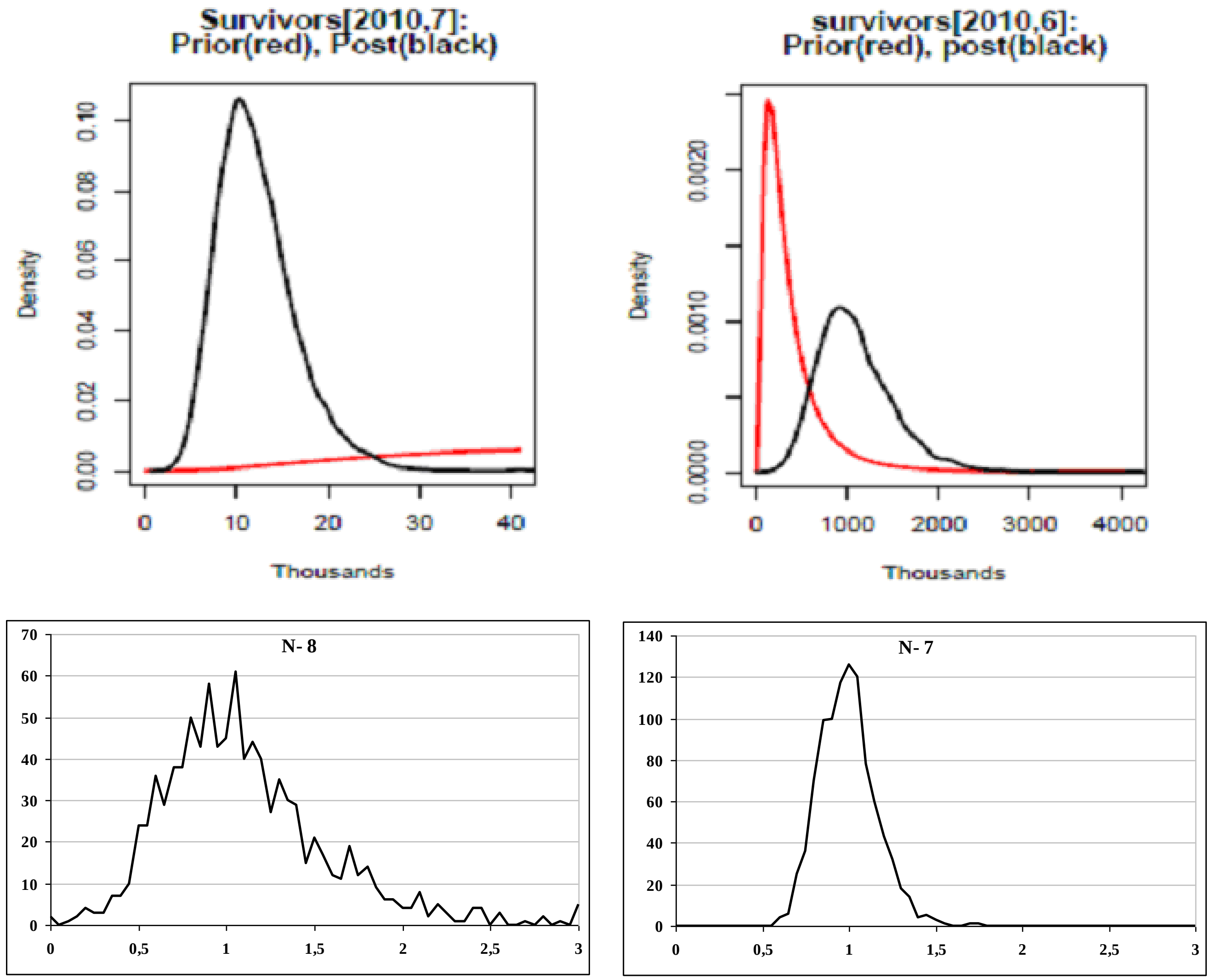


Figure 4 – Posterior distribution of some normalized parameters, as indicated: from Bayesian analysis in the upper squares and from the bootstrap results in the lower ones (Note that Survivors[2010,7] = N[2011,8), which is call here N-8). The Y axis indicates frequency.

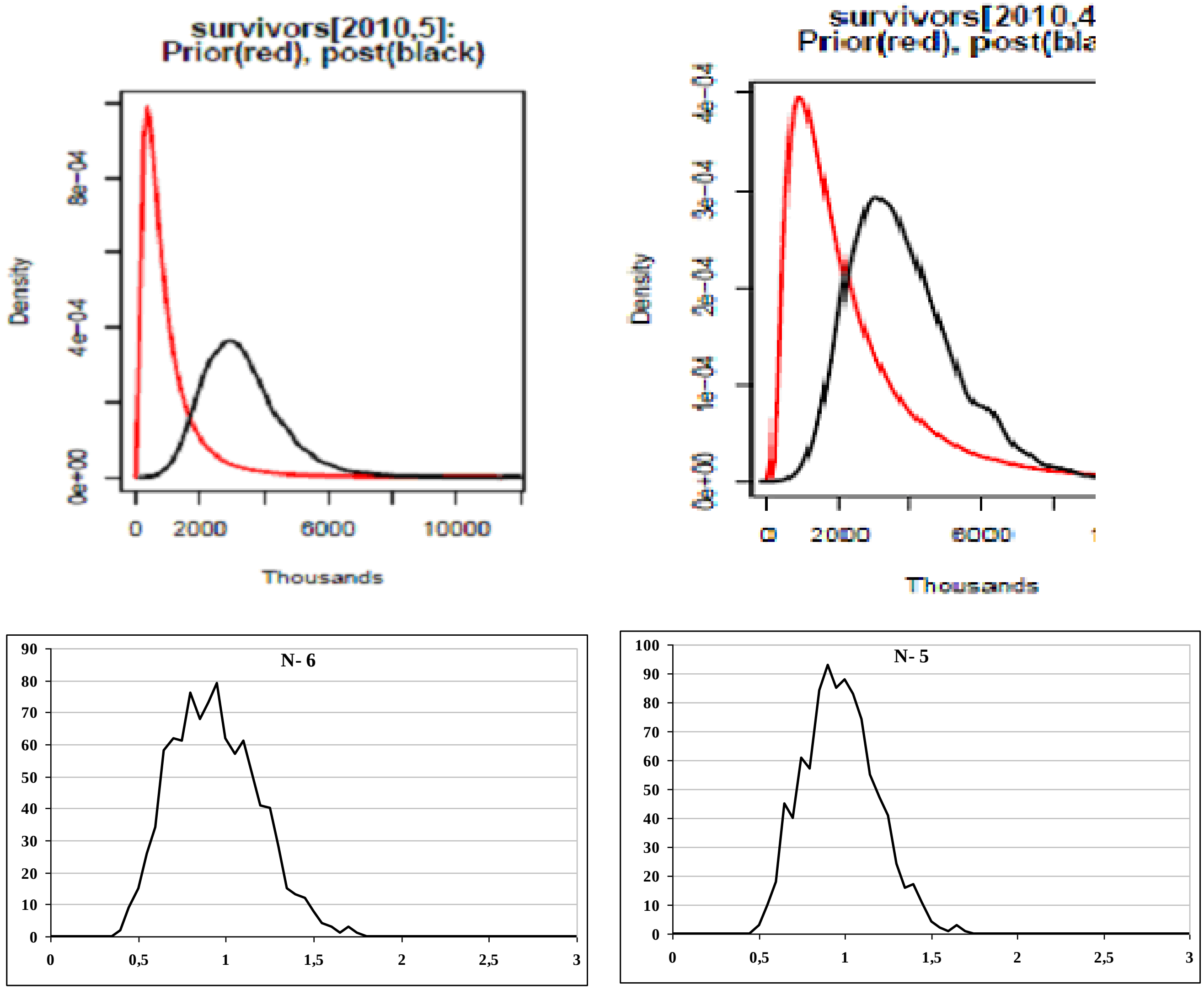


Figure 4 – (continued)

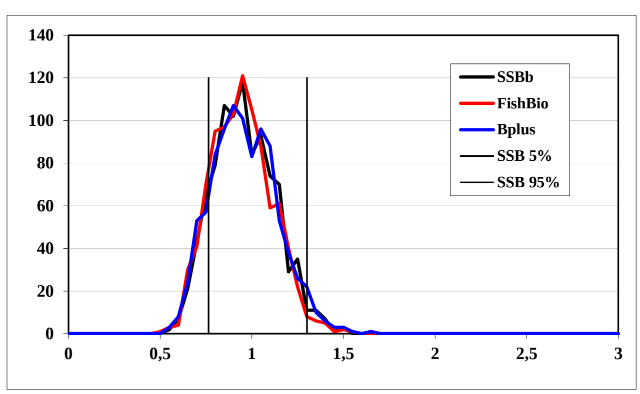


Figure 5 – Distribution of SSB (SSBb), fishable biomass (FishBio) and 5+ biomass (Bpus) of survivors from bootstrap, normalized. The 5% and 95% lines mark the percentiles of SBB in assessment (Bayesian) results. The Y axis indicates frequency.